

Filtrage spatial robuste à partir d'un sous-ensemble optimal d'électrodes en BCI EEG

Alexandre BARACHANT¹, Tetiana AKSENOVA², Stéphane BONNET¹

¹CEA, LETI, DTBS/STD/LE2S,
17 rue des Martyrs, F-38054 Grenoble, France

²RTRA/Fondation Nanosciences
23 rue des Martyrs, 38000 GRENOBLE, France

alexandre.barachant@cea.fr, tetiana.aksenova@cea.fr
stephane.bonnet@cea.fr

Résumé – La réalisation d'une interface cerveau machine EEG nécessite généralement l'utilisation d'un grand nombre d'électrodes, causant la gêne de l'utilisateur et augmentant considérablement le coût calculatoire des traitements. Cependant, un choix judicieux de l'emplacement des ces électrodes peut permettre une réduction importante de leur nombre sans perte significative en performance. Cet article présente une méthode de sélection automatique d'un sous-ensemble quasi optimal d'électrodes et de filtres spatiaux calculés par Common Spatial Pattern (CSP). Cette méthode, basée sur un calcul de coefficient de détermination multiple et l'utilisation du critère d'Akaike, est traitée de manière à résister aux artefacts par l'utilisation d'estimateurs robustes de variance et de matrice de covariance. Il est ainsi montré qu'une réduction très importante du nombre d'électrode est possible sans perte d'information sur les caractéristiques spatiales et que cette méthode résiste parfaitement à un grand nombre d'artefacts lorsque les signaux sont corrompus par des artefacts.

Abstract – EEG-based brain computer interface (BCI) needs a large number of electrodes, leading to user inconvenience and increasing the computational cost of the signal processing. However, a wise choice of location of these electrodes allows a significant reduction in their number without major loss in performance. This paper presents a method for automatic selection of an optimal electrode subset and spatial filter subset (CSP). This is a regression-based method and the use of Akaike criterion provides automation. This methods is also insensitive to outliers using robust estimators of variance and covariance matrix. It is shown that a very substantial reduction in the number of electrodes is possible without loss of significant information and also that this method is completely robust to a large number of outliers.

1 Introduction

Les interfaces cerveau-machine (BCI - Brain Computer Interface) permettent d'établir une communication direct entre un utilisateur et une machine sans recourir à la voie musculaire, ce qui constitue un réel espoir pour les personnes souffrant de graves paralysies [1]. Les BCIs non-invasives utilisent couramment l'électro-encéphalogramme (EEG) comme méthode d'acquisition de l'activité cérébrale. A travers un paradigme, on associe ensuite une ou plusieurs tâches mentales à diverses actions [2]. Chaque tâche mentale possède des caractéristiques fréquentielles, temporelles et spatiales qui lui est propre et que l'on cherche le plus souvent à rehausser par des techniques de filtrage. Les caractéristiques spatiales peuvent être très localisées et plutôt que d'utiliser un grand nombre d'électrodes réparties sur l'ensemble du scalp, il est possible d'utiliser un nombre réduit d'électrodes positionnées judicieusement, ce qui diminuera le coût calculatoire des traitements ainsi que la gêne de l'utilisateur. Cet article décrit une procédure automatique et indépendante de la tâche mentale qui permet de sélectionner un sous-ensemble optimal d'électrodes (sur lesquelles des filtres spatiaux seront calculés). Les signaux

EEG étant très sensibles aux artefacts, principalement de nature musculaire (mouvement, clignement des yeux ...), le problème de robustesse des algorithmes est également abordé.

2 Méthodes

2.1 Sélection d'un sous-ensemble d'électrodes

Procédure : On propose dans cet article une procédure automatique de sélection du sous-ensemble d'électrodes de taille minimale conservant un maximum d'informations sur la spécificité spatiale des tâches mentales. Notre méthode de sélection est basée sur l'emploi de la régression linéaire multiple. La variable à expliquer est la classe $Y_i = \{\pm 1\}$ du test et les variables explicatives sont par convention les variances de chacune des C électrodes. Notre sélection de modèle repose sur l'emploi du coefficient de détermination multiple R^2 qui reflète la force du lien entre les variances des électrodes et la classe du test. Pour pouvoir comparer des modèles avec un nombre variable k de paramètres, i.e. d'électrodes, on utilise dans notre

étude le coefficient de détermination ajusté [3] :

$$R_a^2 = 1 - \frac{N-1}{N-k-1}(1-R^2)$$

avec N le nombre de tests et k le nombre d'électrodes considérées. La procédure proposée se déroule en deux étapes. Dans un premier temps, on recherche pour un nombre donné k d'électrodes, le meilleur sous-ensemble S_k^* de k électrodes parmi les C électrodes qui vérifie :

$$S_k^* = \underset{S_k}{\operatorname{argmax}} R_a^2(S_k)$$

Cette optimisation est faite de manière non exhaustive par une technique modifiée de régression pas à pas ascendante. Dans un second temps, on détermine le nombre minimal d'électrodes k^* qui minimise le critère d'information d'Akaike corrigé [4] sur l'ensemble des sous-ensemble $\{S_k^*\}_{k=1:C}$:

$$k^* = \underset{k}{\operatorname{argmin}} AIC_c(k)$$

avec

$$AIC_c(k) = N \ln(\hat{\sigma}^2(k)) + 2k + \frac{2k(k+1)}{N-k-1}$$

$$\text{avec } \hat{\sigma}^2(k) = \frac{1}{N} \sum_{i=1}^N \varepsilon_i^2(k)$$

où $\hat{\sigma}^2(k)$ est une estimation de la variance des termes résiduels $\varepsilon_i(k)$ de notre modèle bâti sur le sous-ensemble d'électrodes S_k^* . La variance des termes résiduel est lié à $R_a^2(S_k^*)$ par :

$$\hat{\sigma}^2(k) = (1 - R_a^2(S_k^*)) \operatorname{Var}(Y)$$

Variante Robuste : Les estimations de variance seront réalisées de manière avantageuse par des techniques empruntées aux statistiques robustes. En effet, les signaux EEG sont souvent contaminés par des artefacts. Ainsi, la variance des électrodes sera estimée par la déviation absolue médiane (MAD) [5].

$$\operatorname{Var}(\hat{X}) = \left[\frac{1}{0.6745} \operatorname{mediane}(|X - \operatorname{mediane}(X)|) \right]^2$$

2.2 Filtrage spatial par CSP

La méthode de filtrage spatial par Common Spatial Pattern (CSP) est couramment utilisée en BCI [2]. Il s'agit d'effectuer une combinaison linéaire entre les différentes électrodes en vue de maximiser la différence des variances du signal filtré entre les 2 classes [2]. L'algorithme CSP appliqué à nos k^* électrodes fournit un ensemble de filtres spatiaux \mathbf{w}_j , $j \in [1 : k^*]$ et donc de nouveaux signaux \mathbf{z}_j tel que :

$$\mathbf{z}_j = \mathbf{w}_j^T \mathbf{X}$$

avec $\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(t), \dots, \mathbf{x}(T)] \in \mathbb{R}^{k^* \times T}$ la matrice d'un test de T échantillons sur les k^* électrodes retenues. Ces filtres sont obtenus par diagonalisation conjointe des matrices de covariances moyenne de chaque classe [2]. Cependant ces filtres n'ont pas tous un pouvoir discriminant équivalent, certains sont

même totalement inefficaces et peuvent nuire à la qualité finale de la détermination de la classe du test. Il est donc nécessaire de sélectionner un certain nombre de ces filtres. Le pouvoir discriminant de chaque filtre est relié à la valeur propre λ_j qui lui est associé. Une méthode classique de sélection des filtres consiste à trier les $\{|\lambda_j - 0.5|\}_{j=[1:k]}$ et de ne retenir que les M plus grands. Le choix du nombre de filtres M est alors posé arbitrairement ou obtenu grâce à une étude des valeurs propres, ce qui dans les deux cas fait intervenir un spécialiste. De manière analogue à la sélection automatique du sous-ensemble d'électrodes, nous appliquons l'algorithme de sélection sur les signaux filtrés \mathbf{z}_j . En effet, les signaux \mathbf{z}_j peuvent être considérés comme provenant d'électrodes artificielles, il est donc possible d'appliquer la procédure de sélection d'électrodes sur celles-ci. Ainsi sélectionner une électrode artificielle revient à retenir le filtre spatial qui lui est associé. L'aspect robuste de la CSP est traité de manière semblable à ce qui a été décrit précédemment grâce à l'estimation robuste par MCD des matrices de covariances [6].

3 Résultats

Description des données : Afin de tester nos algorithmes, nous avons utilisé le jeu de données IVA de la compétition BCI III [7]. Il s'agit d'enregistrements EEG chez 5 sujets, sur 118 électrodes, pendant deux tâches mentales : l'imagination d'un mouvement de la main droite ou du pied droit. Les jeux de données sont constitués de 280 tests, répartis équitablement entre les deux classes.

Sélection du sous-ensemble d'électrodes et des filtres : la procédure de sélection est appliquée sur les signaux après un filtrage passe-bande entre 10 et 25 Hz. Les figures 1 et 2 nous montrent l'évolution du coefficient de détermination et du critère d'Akaike corrigé en fonction de la taille du sous-ensemble d'électrodes S_k^* (figure 1) ou du sous-ensemble de filtres (figure 2). En pratique la courbe du critère d'Akaike est plutôt plate dans sa partie minimale. Plutôt que de sélectionner le sous-ensemble dont le critère est minimum, il est donc préférable de sélectionner le sous-ensemble de taille minimale dont la valeur du critère a atteint un certain pourcentage de la valeur minimale. Pour l'utilisateur *al*, notre algorithme a sélectionné 8 électrodes et 4 filtres spatiaux.

Spatial Pattern et classification : Le filtrage spatial par CSP nous offre la possibilité de visualiser les principales caractéristiques spatiales des signaux issus de chaque classe [2]. En effet on peut décomposer tout test \mathbf{X} de sorte que $\mathbf{x}(t) = \sum_{j=1}^{k^*} z_j(t) \mathbf{a}_j$ en fonction de caractéristiques temporelles $z_j(t)$ et spatiales \mathbf{a}_j . Les \mathbf{a}_j , appelés aussi Spatial Pattern [2], sont les vecteurs colonnes de la matrice $\mathbf{A} = (\mathbf{W}^{-1})^T$ avec $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{k^*}]$ triés par ordre d'importance. Ils représentent la partie invariante des signaux, dépendant uniquement des tâches mentales. Ils ne doivent donc pas être globalement modifiés

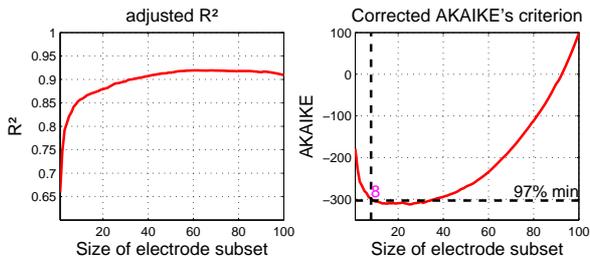


FIGURE 1 – Évolution du coefficient de détermination R_a^2 et du critère d'Akaike en fonction de la taille k du sous-ensemble d'électrodes. Utilisateur *al*

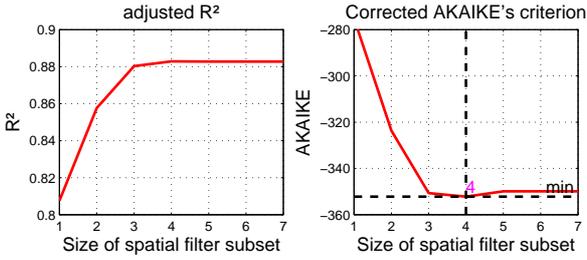


FIGURE 2 – Évolution du coefficient de détermination R_a^2 et du critère d'Akaike en fonction de la taille k du sous-ensemble de filtres spatiaux. Utilisateur *al*

par la sélection d'un sous-ensemble d'électrodes. La figure 3 montre l'une de ces caractéristiques spatiales, qui n'a été que faiblement modifiée par la procédure de sélection. Après filtrage spatial par CSP, il est d'usage de classifier les tâches mentales à partir de la variance des signaux \mathbf{z}_j . La figure 4 montre les résultats de cette classification par LDA en fonction du sous-ensemble d'électrodes S_k^* . Ici on a bien sélectionné le sous-ensemble de taille minimum qui offre les meilleures performances.

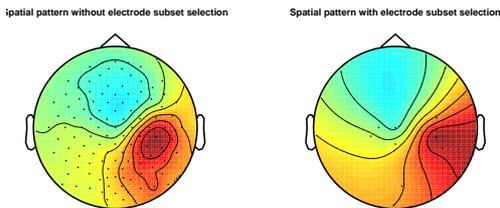


FIGURE 3 – Caractéristique spatiale principale avant (118 électrodes) et après (8 électrodes) réduction du nombre d'électrodes. Utilisateur *al*

Comparaison avec une méthode existante : Il existe plusieurs méthodes permettant la sélection d'un sous-ensemble d'électrodes, cependant aucune ne semble faire office de référence. Une méthode simple consiste à trier par ordre décroissant les électrodes en fonction des coefficients des filtres ($sort|w_{1,i}|$) ou des spatial patterns

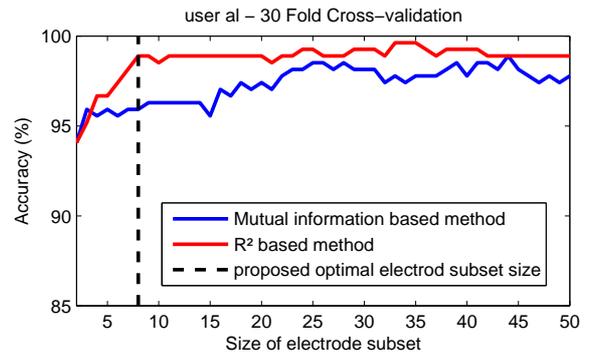


FIGURE 4 – Performance en classification en fonction de la taille du sous-ensemble d'électrodes. Utilisateur *al*

($sort|a_{1,i}|$) [8, 9]. La méthode retenue à titre de comparaison est une méthode basée sur le calcul de l'information mutuelle [10]. La recherche du sous-ensemble optimal se fait aussi par une régression pas à pas ascendante où cette fois on sélectionne le sous-ensemble maximisant l'information mutuelle entre la variance des signaux et la classe des tests. Cette approche en fait une méthode relativement similaire à celle présentée ici, cependant le calcul de l'information mutuelle sur plusieurs dimensions n'est pas trivial et nécessite des étapes de prétraitement (ACI) pour simplifier le calcul. De plus cette méthode ne permet pas de retenir automatiquement le nombre d'électrode optimal. Il apparaît sur les résultats donnés en figure 4 que la méthode décrite dans cet article arrive plus rapidement au sous-ensemble optimal et que les deux méthodes finissent par converger. La table 1 contient les résultats en classification pour un nombre donné d'électrodes et de filtres. Les deux méthodes y sont comparées ainsi que la méthode habituelle consistant à prendre l'ensemble des électrodes et des filtres disponibles. Il en ressort un net avantage pour notre méthode. De plus on observe une perte de performance importante due à un sur-apprentissage lorsque l'on utilise l'ensemble des 118 filtres. De même lorsqu'on pose un nombre arbitraire de filtres, ici 10, les performances ne sont pas optimales. L'utilisation d'une méthode de sélection automatique prend donc tout son sens.

Robustesse : La robustesse des procédures est testée et validée via l'addition d'artefacts artificiels dans les signaux du jeu de données. Ces artefacts suivent la loi de probabilité suivante :

$$\xi \sim (1 - \varepsilon)\delta_0 + \varepsilon\mathcal{U}([- \mu, \mu]) \quad (1)$$

Où $\mathcal{U}([- \mu, \mu])$ représente une distribution uniforme entre $-\mu$ et μ , δ_0 la distribution de Dirac en 0. ε et μ sont des paramètres de simulation, $\varepsilon \leq 1$ représente la probabilité d'apparition des artefacts, $\mu = \kappa\hat{\sigma}$ l'amplitude de ces artefacts, fixée en fonction de l'écart type des signaux de chaque électrode $\hat{\sigma}$ multiplié par un facteur κ . Pour les résultats présentés, on se place volontairement dans un cas où les artefacts sont suffisamment nombreux et de forte amplitude pour se placer au delà du point de rupture des méthodes non-robustes ($\varepsilon = 0.1, \kappa = 30$). Dans un premier temps, la robustesse de l'algorithme de sélection du

TABLE 1 – Comparaison des performance en classification pour les différentes méthodes (R^2 : méthode basée sur le coefficient de détermination, MI : Information mutuelle, All : pas de sélection). 50-fold Cross Validation

<i>User</i>	Méthodes	N Elec	N Filter	Classif (%)
<i>al</i>	R^2	8	4	98.8
	MI	8	4	97.2
	All	118	118 10	97.2 98
<i>aa</i>	R^2	17	6	83.2
	MI	17	6	83.2
	All	118	118 10	64.4 75.5
<i>av</i>	R^2	8	5	70
	MI	8	5	68
	All	118	118 10	57.2 65.5
<i>aw</i>	R^2	20	7	96.4
	MI	20	7	91.6
	All	118	118 10	84.8 94.4
<i>ay</i>	R^2	19	7	88.4
	MI	19	7	87.6
	All	118	118 10	74.8 82.4
<i>moyenne</i>	R^2	14.4	5.8	87.3
	MI	14.4	5.8	85.5
	All	118	118 10	75.7 83.2

sous-ensemble d'électrodes est testée indépendamment de la CSP. La table 2 nous fournit les 4 premières électrodes sélectionnées pour l'utilisateur *al*. Dans cet exemple il apparait que la méthode robuste est insensible aux artefacts. En revanche on s'aperçoit que la méthode robuste ne sélectionne pas exactement le même sous-ensemble que la méthode classique. En effet, différentes électrodes peuvent apporter le même contenu informationnel, le choix de l'électrode à retenir parmi ces électrodes se fait sur la base d'une très faible différence. De plus les électrodes sélectionnées sont cohérentes avec les tâches mentales, en effet C_3 et C_{PZ} sont respectivement situées au dessus des zones dédiées à la main droite et au pied. Dans un deuxième temps, on évalue la robustesse de la chaîne complète avec classification. Les résultats sont reportés en table 3. Sur le jeu de données contenant des artefacts, la méthode non-robuste est incapable d'obtenir un résultat correct, en revanche la méthode robuste affiche des résultats similaires a ceux obtenus sur le jeux de données sans artefacts.

4 Conclusion

La méthode présentée dans cet article permet de réduire avantageusement le nombre d'électrodes utilisées lors d'expérimentations BCI, diminuant ainsi grandement le temps de préparation et la puissance de calcul nécessaire. Cette méthode étant entièrement automatique et indépendante des tâches mentales utilisées, son application ne nécessite pas l'intervention d'un

TABLE 2 – 4 premières électrodes sélectionnées pour l'utilisateur *al* sur signaux avec et sans artefacts. $\varepsilon = 0.1, \kappa = 30$

<i>Electrodes</i>	Méthode classique	Méthode robuste
Sans artefacts	$C_3 - C_{PZ}$	$C_3 - C_{PZ}$
	$CC_{P6} - FF_{C4}$	$CC_{P6} - TP_8$
Avec artefacts	$CF_{C6} - PP_{O7}$	$C_3 - C_{PZ}$
	$PP_{O1} - C_6$	$CC_{P6} - TP_8$

TABLE 3 – Performance en classification pour les méthodes robustes/non-robustes avec et sans artefacts, $\varepsilon = 0.1, \kappa = 30$. 50-Fold Cross Validation.

<i>User</i>	Méthode classique		Méthode robuste	
	Artefacts non	oui	non	oui
<i>al</i>	98.8	53.2	96	94.4
<i>aa</i>	83.2	55.6	81.6	83.2
<i>av</i>	70	53.6	69.7	72
<i>aw</i>	96.4	53.2	94	93.6
<i>ay</i>	88.4	59.6	94.4	94.4
<i>moyenne</i>	87.3	55.1	87.2	87.5

spécialiste. De plus, la robustesse des techniques utilisées permet de libérer l'utilisateur d'un grand nombre de contraintes et autorise l'emploi de cette approche dans le cadre d'une situation de vie courante.

Références

- [1] Mikhail A Lebedev and Miguel A L Nicolelis. Brain-machine interfaces : past, present and future. *Trends in Neurosciences*, 29, September 2006.
- [2] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller. Optimizing spatial filters for robust EEG Single-Trial analysis. *Signal Processing Magazine*, 2008.
- [3] M. Tenenhaus. *Statistique : Méthodes pour décrire, expliquer et prévoir*. Dunod, December 2006.
- [4] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 1974.
- [5] Xinyi Yong, Rabab K. Ward, and Gary E. Birch. Robust common spatial patterns for EEG signal preprocessing. In *Engineering in Medicine and Biology Society*, 2008.
- [6] T Aksenova, A Barachant, and S Bonnet. Spatial filtering optimisation in motor imagery EEG-based BCI. In *Neurocomp 2008*.
- [7] Data set IVa for the BCI competition III. http://ida.first.fhg.de/projects/bci/competition_iii/desc_IVa.html.
- [8] Yijun Wang, Shangkai Gao, and Xiaorong Gao. Common spatial pattern method for channel selection in motor imagery based brain-computer interface. In *Engineering in Medicine and Biology Society*, 2005.
- [9] Xiaorong Gao Bin Lou, Bo Hong and Shangkai Gao. Bipolar electrode selection for a motor imagery based brain-computer interface, September 2008.
- [10] Tian Lan, D. Erdogmus, A. Adami, M. Pavel, and S. Mathan. Salient EEG channel selection in brain computer interfaces by mutual information maximization. In *Engineering in Medicine and Biology Society*, 2005.